PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : C12N 15/86, A61K 48/00 — A1

(11) International Publication Number: WO 97/07223

(43) International Publication Date: 27 February 1997 (27.02.97)

(21) International Application Number: PCT/EP96/00761

(22) International Filing Date: 23 February 1996 (23.02.96)

(30) Priority Data:
195 30 412.8 18 August 1995 (18.08.95) DE

(71)(72) Applicants and Inventors: VON MELCHNER, Harald [DE/DE]; Abt. Hämatologie, Universitätsklinikum Frankfurt, Theodor-Stern Kai 7, D-60590 Frankfurt am Main (DE). RUSS, Andreas, Peter [DE/DE]; Abt. Hämatologie, Universitätsklinikum Frankfurt, Theodor-Stern-Kai 7, D-60590 Frankfurt am Main (DE). GREZ, Manuel [DE/DE]; Paul-Ehrlich-Strasse 42, D-60596 Frankfurt am Main (DE).

(74) Agent: GRÜNECKER, KINKELDEY, STOCKMAIR & SCHWANHÄUSSER; Maximilianstrasse 58, D-80538 München (DE).

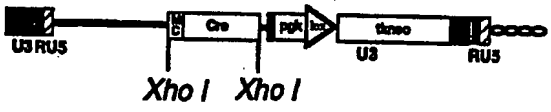
(81) Designated States: AL, AM, AT, AU, AZ, BB, BG, BR, BY, CA, CH, CN, CZ, DE, DK, EE, ES, FI, GB, GE, HU, IS, JP, KE, KG, KP, KR, KZ, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, TJ, TM, TR, TT, UA, UG, US, UZ, VN, ARIPO patent (KE, LS, MW, SD, SZ, UG), Eurasian patent (AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).

Published
With international search report.

(54) Title: SELF-DELETING VECTORS FOR GENE THERAPY

U3pgkIxtkneo

U3pgkIxtkneoMCCre



(57) Abstract

The invention involves the development of vectors for somatic gene therapy. The vectors transduce a complete transcriptional unit containing a promoter, a protein coding sequence and a polyadenylation sequence into the genome of mammalian cells. Upon integration, the vectors delete most viral and non-viral sequences unrelated to transcriptional unit thus avoiding common problems encountered with conventional retrovirus vectors such as repression of gene expression by transcriptional silencing, mobilization of endogenous retroviruses, activation of oncogenes or development of an immune response. The invention exploits (i) the natural life cycle of retroviruses, involving duplication of the terminal control regions U5 and U3 to generate long terminal repeats (LTR) and (ii) the ability of site specific recombinases to excise any sequences positioned between two specific target sequences from the mammalian genome. Thus, the retroviruses of the invention transduce the coding sequences of a site specific recombinase and at least one recombinase-specific target sequence into the genome along with the transcriptional unit expressing a therapy gene.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AM	Armenia	**GB**	United Kingdom	**MW**	Malawi
AT	Austria	**GE**	Georgia	**MX**	Mexico
AU	Australia	**GN**	Guinea	**NE**	Niger
BB	Barbados	**GR**	Greece	**NL**	Netherlands
BE	Belgium	**HU**	Hungary	**NO**	Norway
BF	Burkina Faso	**IE**	Ireland	**NZ**	New Zealand
BG	Bulgaria	**IT**	Italy	**PL**	Poland
BJ	Benin	**JP**	Japan	**PT**	Portugal
BR	Brazil	**KE**	Kenya	**RO**	Romania
BY	Belarus	**KG**	Kyrgystan	**RU**	Russian Federation
CA	Canada	**KP**	Democratic People's Republic of Korea	**SD**	Sudan
CF	Central African Republic			**SE**	Sweden
CG	Congo	**KR**	Republic of Korea	**SG**	Singapore
CH	Switzerland	**KZ**	Kazakhstan	**SI**	Slovenia
CI	Côte d'Ivoire	**LI**	Liechtenstein	**SK**	Slovakia
CM	Cameroon	**LK**	Sri Lanka	**SN**	Senegal
CN	China	**LR**	Liberia	**SZ**	Swaziland
CS	Czechoslovakia	**LT**	Lithuania	**TD**	Chad
CZ	Czech Republic	**LU**	Luxembourg	**TG**	Togo
DE	Germany	**LV**	Latvia	**TJ**	Tajikistan
DK	Denmark	**MC**	Monaco	**TT**	Trinidad and Tobago
EE	Estonia	**MD**	Republic of Moldova	**UA**	Ukraine
ES	Spain	**MG**	Madagascar	**UG**	Uganda
FI	Finland	**ML**	Mali	**US**	United States of America
FR	France	**MN**	Mongolia	**UZ**	Uzbekistan
GA	Gabon	**MR**	Mauritania	**VN**	Viet Nam

SELF-DELETING VECTORS FOR GENE THERAPY

FIELD OF INVENTION

This invention relates to specially constructed vectors which transduce transcriptional units into eukaryotic cells such that preferably all sequences unrelated to the transcriptional unit are eliminated upon integration. Only one transcriptional unit comprising any natural or synthetic promoter/enhancer sequence, protein coding sequence and polyadenylation site is retained in the genome. By eliminating themselves from the genome the vectors of the invention circumvent the problems encountered with conventional vectors, e.g. retroviruses and vectors thereof, such as transcriptional interference with transduced genes or genes adjacent to the integration site, activation of cellular oncogenes, mobilization of endogenous retroviruses and development of an immune response.

BACKGROUND OF THE INVENTION

Retroviruses are RNA viruses that replicate through a DNA intermediate. Flanking the ends of the viral RNA genome are short sequence repeats (R) and unique sequences (U5 and U3) that control DNA synthesis, integration, transcription, and RNA processing. Between the control regions are coding sequences for the major structural proteins of the virus particle (gag and env) and for enzymes found in particles (pol, protease, reverse transcriptase and integrase) (Figure 1).

Shortly after infection, viral RNA is converted into DNA by reverse transcriptase. The process is initiated by cellular tRNA which by binding to a complementary region within the viral genome serves as elongation primer. This region, also

termed - primer binding site- (PBS) is located immediately downstream of U5 and is essential for virus replication. Prior to integration, terminal sequences of the viral genome are duplicated such that the retroviral genome is flanked by long terminal repeats (LTRs) each containing the U3, R and U5 regions. Linear DNA molecules of this type integrate into the genome (Figure 1).

LTR sequences are maintained in the integrated retrovirus, also termed -provirus-, except that two nucleotides (nt) are lost from each end. Cellular DNA sequences also are unaltered except that upon integration, 4-6 nt are duplicated such that the provirus is flanked at each end by 4-6 nt repeats. As a provirus, the retroviral genome is replicated with cellular DNA and transcribed as a cellular gene by RNA polymerase II. Provirus transcription is controlled by promoter/enhancer sequences located in the U3 region of the 5'LTR. Polyadenylated transcripts initiate at the junction between U3 and R (cap site) in the 5'LTR and terminate in R of the 3'LTR that contains the signal for polyadenylation. Full-length (genomic) RNA is transported from the nucleus to the cytoplasm and either packaged into virus particles that bud from the cell or are translated to yield gag and pol proteins. A fraction of the RNA is spliced to yield mRNA encoding env.

It is possible to adapt retroviruses to transduce genes into mammalian genomes. Provided that certain control sequences within the LTRs remain unaltered, the retroviral genome can be deleted without impairing its ability to replicate in cells that express proteins necessary for reverse transcription, integration and particle formation. For this, vector DNA is transfected into cell lines that contain complete retroviral genomes or helper viruses. The helper viruses are constructed such that they cannot assemble into particles, due to a small deletion encompassing a sequence (Y) between U5 and gag. Since the vector DNA does not

contain the Ψ deletion, recombinant transcripts are packaged and expelled from the cells as virus particles. In addition to Ψ, gag sequences can also enhance the ability of the vectors to be packaged.

To date, retroviral vectors are the most efficient means to transduce foreign genes into mammalian cells. Accordingly, retroviruses are used in over 80% of all approved gene therapy trials. However, several factors undermine the practical use of conventional retroviruses as gene therapy vectors. First, transduced genes are often inactivated by methylation or binding of transcriptional repressors to the viral genome (1-3). Since these repressors were shown to bind the primer binding site of several retroviruses, its simple deletion would preclude virus replication. Second, since retroviruses integrate mostly randomly throughout, the genome, integrations sometimes result in mutations that augment the expression of adjacent genes (4, 5). Activation of adjacent proto-oncogenes followed by malignant transformation of the infected cells has been described (6). The activation mechanism involves transcriptional enhancement either by upstream U3 promoters or nearby U3 enhancers. Third, viral vector sequences such as packaging signals and leader sequences can potentially recombine with endogenous retroviruses yielding new and unpredictable forms of infectious virus (7-9). Finally, viral and non-viral sequences of some vectors may trigger an immune response which eliminates the transduced cells (10). A review on gene therapy is given in (26).

The object underlying the present invention is to provide novel vectors which do not show the disadvantages of prior art vectors.

Said object is achieved by a vector system useful for gene therapy comprising at least one coding sequence for a site-

specific recombinase and at least one target sequence being specifically recognized by said recombinase.

The vector system can be a DNA or RNA vector. Presently, there are several vectors available which can be used for gene therapy and which are suitable as a starting material for the construction of a vector system according to the invention.
The site-specific recombinase can be any recombinase which specifically recognizes a target sequence. The prior art provides several examples for site-specific recombinases and the corresponding target sequence. The location of the sequence coding for the recombinase and the target sequence within the vector can be chosen by the skilled person depending on which parts of the vector shall be deleted after being incorporated into the mammalian genome. Optionally, the site-specific recombinase can also be encoded by a separate vector which is simultaneously contained in the cell with the vector containing the target sequence and the further sequences required for incorporating said vector into the genome. In a further alternative the recombinase can also be added as a protein to the cell containing the vector with the target sequence.

In a preferred embodiment the combination of recombinase and target sequence used is the recombinase Cre and the target sequence loxP.

In a further preferred embodiment the recombinase is Flp and the target sequence is frt.

The vector according to the present invention may comprise any transcriptional unit, which transcriptional unit comprises a gene coding for the desired function to be introduced into the mammalian genome. In many cases said gene will code for a protein which is not properly prepared

by the cell which receives the vector according to the invention.

In a further preferred embodiment the vector according to the invention is a retroviral vector, most preferably a retroviral DNA.

In case that a retroviral vector is used the target sequence is preferably inserted into the U3 and/or U5 region, which region may additionally contain the transcriptional unit. In a further preferred embodiment the vector according to the invention may also comprise a viral promoter and/or enhancer.

As the result of the introduction of the vector according to the invention into a mammalian cell a cell is obtained which, after deletion of the non-desired vector parts from the genome, contains at least one target sequence, which was introduced by the vector. In most cases where the vector also comprised the additional transcriptional unit said additional transcriptional unit will be contained in the mammalian cell genome. In the event that the transcriptional unit contains a gene which normally is also contained in the mammalian cell but for some reason non-functional then the gene introduced into said cell via the vector according to the invention will have a different chromosomal environment compared to the gene as it naturally occurs in the cell. When the vector sequences, e.g. the proviral genome, is deleted from the chromosome in general a few nucleotides of the proviral genome remain in the genome of the cell. "Essentially no proviral genome" as used herein, therefore, means that preferably less than 600 bp nucleotides, more preferably less than 100 bp stay in the genome.

By using the vector system according to the present invention any DNA can be incorporated into the genome of a

mammalian cell. The location of the target sequence within the vector allows to predetermine which parts of the vector shall be deleted from the genome after incorporation of the vector.

The invention involves the development of novel vectors, particularly retroviruses, to transduce transcriptional units for therapeutic or non-therapeutic purposes into the genome of eukaryotic cells. The vectors are equipped with a site-specific recombination system including a recombinase such as Cre or Flp and at least one specific target sequence such as loxP or frt. The system is activated upon integration such that all viral and non-viral sequences unrelated to the transcriptional unit are eliminated. Retroviruses of this type preferably circumvent all unwanted side effects encountered with conventional retrovirus vectors.

The preferred retroviruses of the invention contain a transcriptional unit including a promoter, protein coding sequence and polyadenylation sequence preferably in the U3 or U5 regions. Also preferably in the U3 or U5 regions, the retroviruses of the invention contain at least one synthetic or natural target sequence of a site specific recombination system. In such a system, DNA fragments flanked by target sequences that have the same orientation are eliminated by the corresponding recombinase (Figure 2).

The retroviruses of the invention duplicate the target sequences inserted into the U3 or U5 regions during replication. This positions most of the viral genome between identical target sequences and enables the recombinase to delete the bulk of the provirus. Only one copy of the transduced transcriptional unit that contains the gene of interest remains in the genome. The expression of the recombinase is either achieved in *trans*, by introducing the protein or a recombinase-expressing plasmid

into the transduced cells or, preferably in *cis* by expressing the recombinase from the provirus itself (Figure 3,4). In this case, the recombinase coding sequences are preferably placed between the target sequences outside of the proviral control regions (LTRs). Their expression is preferably controlled by a second synthetic or natural promoter. A polyadenylation signal is preferably provided by the R-region of the 3'LTR. In the event of an inverse orientation of the recombinase cassette relative to the provirus, transcripts may terminate in cryptic proviral polyadenylation signals or in an additional synthetic or natural polyadenylation sequence cloned into the virus in inverse orientation.

The retroviruses of the invention are preferably enhancer and/or promoterless. However, the retroviruses of the invention may continue to contain their own promoter and enhancer sequences.

The retroviruses of the invention are preferably used to transduce therapy genes into mammalian cells, however, they may also be used for basic research.

The retroviruses of the invention may also transduce potentially hazardous sequences as long as these are deleted in due time following integration. For example, a potential application would be vectors based on human retroviruses (HIV, HTLVI).

The invention includes mammalian cells containing at least one site specific recombination target (e.g. loxP or frt) and at least one protein coding sequence transduced by the retroviruses of the invention. These cells are largely devoid of retroviral sequences.

Finally, the invention includes procedures for transducing cDNA sequences into the mammalian genome such as

transfection of producer cells with the retroviruses of the invention, infection of mammalian cells and expression of site specific recombinases associated with deletion of the recombinase-expressing cassettes.

BRIEF DESCRIPTION OF THE FIGURES

Figure 1 is a schematic diagram or the genome of prior art retroviruses.

Figure 2 is a schematic diagram of a site specific recombination system.

Figure 3 and 4 is a schematic diagram of the retroviruses of the invention where P = promoter; Ts = site specific target sequence; PCS = protein coding sequence, Rec = recombinase.

Figure 5 is a schematic diagram of the retrovirus vectors U3pgk1xtkneo and U3pgk1xtkneoMCCre.

Figure 6 shows a Southern blot analysis of U3pgk1xtkneo expressing clones before and after transfecting MCCre. In the first 6 lanes the DNA was digested with the enzymes NdeI and XbaI that do not cut within the provirus. Lanes 7-12 contain DNAs digested with HindIII from Cre-expressing (+) and non-expressing (-) clones.

Figure 7 shows a Southern blot analysis with DNAs derived from U3pgk1xtkneo (A) and U3pgk1xtkneoMCCre (B) expressing clones. Note the decrease in signal intensity of constant HindIII fragments from Cre-expressing clones (B). Variable bands correspond to the number of proviruses.

Figure 8 is a schematic diagram of the retrovirus vectors pggSVCreU31xpgkpuro and pggSVCreU31xSVpuro.

Figure 9 shows the recombination of U3lxpgkpuroSVCre proviruses in NIH3T3 cells. (A) Predicted structure of proviruses before and after site-specific recombination. (B) Southern blot analysis of U3lxpgkpuroSVCre proviruses. Genomic DNAs were cleaved with EcoRI, fractionated on agarose gels, blotted to nylon filters, and hybridized to a ^{32}P-labeled pgk- probe. Lanes 1-14, U3lxpgkpuroSVCre expressing clones 1-14. The ubiquitous 7 kb constant band represents the endogeous pgk promoter. (C) PCR analysis of U3lxpgkpuroSVCre expressing clones. Genomic DNA was amplified using Cre- (top) or b-actin- (bottom) specific primers. Amplification products were resolved in 1% agarose gels and visualized by Ethidium-bromide staining in lanes as follows: M, molecular weight standards (1 kb BRL ladder); 1-14, U3lxpgkpuroSVCre expressing clones 1-14; 15, single copy U3pgk1xtkneo expressing clone (negative control); 16, single copy U3pgk1xtkneoMCCre expressing clone (positive control). **Figure 10** shows the recombination of U3lxSVpuroSVCre proviruses in NIH3T3 cells. (A) Predicted structure of proviruses before and after site-specific recombination. (B) Southern blot analysis of U3lxSVpuroSVCre proviruses. Cell DNAs were cleaved with EcoRI (left) or Hind III (right), processed as described in the legend to Figure 2, and hybridized to a ^{32}P-labeled SV40-probe. Lanes 1-12, U3lxSVpuroSVCre expressing clones 1-12. (C) PCR analysis of U3lxSVpuroSVCre expressing clones. Genomic DNA was amplified using Cre- (left) or b-actin- (right) specific primers. Amplification products were resolved in 1% agarose gels an visualized by Ethidium-bromide staining in lanes as follows: M, molecular weight standards (BRL 1 kb ladder); 1-12, U3lxSVpuroSVCre expressing clones 1-12; 13-14, single copy U3pgk1xtkneo expressing clones (positive controls); 15-16, single copy U3pgk1xtkneoMCCre expressing clones (negative controls).

The term "vector system" means at least one vector which can introduce into the genome of a cell a coding sequence for a

recombinase and a target sequence for said recombinase. The recombinase encoding sequence and the target sequence can be located on two different vectors but are preferably located in one vector. Optionally, the recombinase can be introduced as a protein into the cell, so that the vector system needs not to comprise a coding sequence for the recombinase.

The term "retrovirus" refers to any RNA virus that replicates through a DNA intermediate. Such viruses can include those that require the presence of other viruses, such as helper viruses, to be passaged. Thus, retroviruses are intended to include those containing substantial deletions or mutations in their RNA.

The term "control region" refers to that region of a retrovirus that is duplicated after infection and prior to integration. Control regions include U3 and U5 regions. Such regions also include LTR regions.
The term "transcriptional unit" refers to a sequence of nucleic acids that includes a natural or synthetic promoter, a protein coding sequence and a polyadenylation signal. Promoters can include an enhancer.

The term "protein coding sequence" means a nucleotide sequence encoding a polypeptide chain that has a therapeutic value or interferes with the cellular metabolism in some way. It also includes polypeptides which can be used to distinguish cells expressing the polypeptide chain from cells not expressing the polypeptide chain, commonly referred to as "selectable markers".

The term "target sequence" or "site specific target sequence" refers to synthetic or natural nucleotide sequences that are recognized by a site-specific recombinase. Examples of such sequences are loxP (11-13)

derived from P1 phage or frt derived from S. cerevisiae (14-16).

The term "recombinase" or "site specific recombinase" refers to a synthetic, natural or recombinant enzyme that binds, cleaves and recombines specific target sequences. Examples for such enzymes are Cre-recombinase from P1 phage (13) or Flp-recombinase from S. cerevisiae (14).

The present invention involves self-deleting vectors, preferably retroviruses, that are preferably used to transduce genes of therapeutic value into somatic cells.

Figure 5 shows two preferred embodiments of the invention. In the upper vector -pggU3pgklxtkneo- a transcriptional unit consisting of a murine phosphoglycerate kinase (pgk) promoter (17), a loxP target sequence (11) and a thymidine-kinase/neomycinphosphotransferase (tkneo) fusion gene was inserted into the 3'-U3 region of an enhancerless MoMuLV retrovirus vector. In the lower vector an additional transcriptional unit encoding MCCre (where Cre = Cre recombinase, MC = HSV thymidine kinase promoter fused to a Polyoma large T enhancer) was inserted between the LTRs into the body of the virus. Virus replication and LTR-mediated duplication places MCCre along with other viral and non-viral sequences between loxP sites enabling Cre recombinase to delete most of the integrated provirus except for one copy of a pgk-lx-tkneo containing LTR.

Figure 8 shows two additional preferred embodiments of the invention. In both constructs the MC promoter was replaced by the simian virus 40 (SV40) promoter and the tkneo fusion gene by the puromycin resistance gene. In the upper vector - pggSVCreU3lxpgkpuro- the puromycin resistance gene is controlled by a pgk promoter. In the lower vector - pggSVCreU3lxSVpuro the same U3 gene is controlled by an SV40

promoter. In these examples the loxP site in the 3'LTR was placed upstream of the transcriptional unit.

Although in the above preferred embodiments the transcriptional unit and the site specific recombination target are placed within U3, this is not obligatory. Both transcriptional unit and target sequences can also reside in the U5 region, or other parts of the vector, depending on which parts of the vector shall be deleted.

Finally, Cre recombinase can also be provided in *trans* either as a native protein or as a recombinase expressing plasmid.

EXAMPLE 1

To investigate whether provirus sequences flanked by site specific recombination targets can be deleted following integration, we have inserted the transcriptional unit -pgk-loxP-tkneo into the U3 region of an enhancerless Moloney murine leukemia virus (Figure 5).

Plasmids

The sequences for Cre recombinase and loxP were derived from pMCCre and pGEM30, respectively (11) The mouse phosphoglycerate-kinase-promoter (pgk) was obtained from ppgkCat (18) and the SV40/puromycin-acetyltransferase cassette from pBABEpuro (19). The tkneo gene was obtained by ligating an NheI/SpeI PCR amplification product of the HSV thymidine kinase coding sequence (20) to an SpeI/NheI amplification product of the neomycin-phosphotransferase coding sequence (21). In frame fusion was achieved by deleting the thymidine kinase stop codon and the Neo- AUG. A pgklxtkneo cassette was cloned blunt ended into the unique NheI site of pggU3en(-) to obtain pggU3pgklxtkneo. pggU3en(-

) was derived from pggU3Neoen(-) (21) by deleting neo and subcloning the viral sequences as an SstI fragment into the backbone of pBABEpuro. pgklxtkneo was assembled in pBluescriptIIKS (Stratagene) by inserting loxP as a PstI/EcoRI fragment of pGEM30, pgk as an XbaI/BglII fragment of pgkCAT and tkneo as a blunt ended fragment into the corresponding sites of the pBluescriptIIKS polylinker. To obtain pggU3pgklxtkneoMCCre, a PCR amplification product of MCCre was cloned into the unique XhoI site of pggU3pgklxtkneo.

Cells and viruses

NIH3T3 and BOSC23 (22) cells were grown in DMEM (Gibco) medium supplemented with 10% fetal bovine serum (Gibco). Helper virus free recombinant retroviruses were obtained by transient transfection of BOSC23 cells as described by Pear et al. (22). Infections were performed by incubating for 24 hours 10^5 NIH3T3 cells with filtered viral supernatants in the presence of 4 µg/ml polybrene (Aldrich). Provirus expressing clones were isolated by selecting for 7 days in medium containing 2 µg/ml puromycin (Sigma) or 1 mg/ml G418 (Gibco).

Southern blot hybridization

DNA hybridizations were performed with ^{32}P-labeled probes as previously described (21). Southern blots were scanned with a PhosphoImager (Molecular Dynamics) and analyzed with ImageQuantNT software (Molecular Dynamics).

Results

Previous investigations have shown that the U3 region of MoMuLV can tolerate relatively large amounts of extra-sequences (up to 5 kb, (23)). These are duplicated along

with viral sequences to generate long terminal repeats (LTRs) that flank the integrated provirus (reviewed in (24)). Therefore, we predicted that following insertion of the target sequence for site-specific recombination -loxP- into the 3'-U3 region of a retrovirus vector, LTR-mediated duplication would place the viral genome between two loxP sites, thus rendering it susceptible to excision by Cre-recombinase.

To investigate this prediction, a transcriptional unit consisting of a murine phosphoglycerate kinase (pgk) promoter (17) a loxP target sequence (11), and a thymidine-kinase/neomycinphosphotransferase (tkneo) fusion gene was inserted into the 3'-U3 region of an enhancerless MoMuLV retrovirus vector (pggU3en(-), (21)) to obtain -pggU3pgklxtkneo- (Figure 5). This plasmid was transfected into BOSC23 helper cells (22) to produce recombinant virus that was used to infect NIH3T3 cells. LTR-mediated duplication should generate in the infected cells proviruses flanked by pgklxtkneo (Figure 6). To test this, twelve neomycin-resistant clones obtained, after selecting in G418, were analyzed by Southern blotting. LTR-mediated duplication was confirmed in each case by using restriction endonucleases that cleave the LTRs (data not shown). To determine whether Cre-recombinase would delete the sequences flanked by loxP and thus fuse the 5'pgk promoter to the 3'tkneo gene (Figure 6), two cell lines expressing U3pgklxtkneo were co-transfected with the expression plasmids MCCre (11) and pBABEpuro (19) encoding for Cre-recombinase and puromycin-resistance, respectively. Genomic DNAs extracted from cell pools obtained by selecting in puromycin were digested with HindIII, an endonuclease that cleaves the LTRs (Figure 6). When hybridized to a neo-probe, non-recombined proviruses generate a constant band of 4.7 kb which accommodates the sequences flanked by loxP (Figure 6, lanes 7-12). This band was significantly fainter or disappeared completely in clones expressing Cre (Figure

6, lanes 8, 9, 11, 12), indicating that most of the proviruses have recombined. Thus, loxP sites inserted into U3 enable Cre recombinase to excise most of the provirus except for a single LTR. The part of proviral DNA left in the chromosome after deletion of most of proviral DNA can be easily examined by sequence analysis of the chromosomal insertion site and sequence comparison with the vector used for introducing the foreign gene.

EXAMPLE 2

Since we sought to develop a vector that carries all required elements for self-deletion into mammalian cells, we investigated whether loxP sites inserted into U3 undergo recombination when Cre is expressed from the same provirus. Coding sequences for Cre controlled by an MC promoter were inserted between the LTRs into pggU3pgklxtkneo to obtain pggU3pgklxtkneoMCCre. To obtain infectious virus, pggU3pgklxtkneoMCCre was transfected into BOSC23 cells as described in example 1. Neomycin resistant clones obtained after infecting NIH3T3 cells and selecting in G418 were analyzed by Southern blotting as described in example 1. Proviruses derived from this vector were expected to excise themselves after integration. Genomic DNA from several independent U3pgklxtkneo and U3pgklxtkneoMCCre expressing clones was digested with HindIII, a restriction endonuclease that cleaves both proviruses downstream of tkneo. When hybridized to neo, non-recombined proviruses maintain all sequences flanked by loxP and as a result, generate a constant band of 4.7 and 6 kb, respectively. Although this band was still present in Cre expressing clones, it was 5-10 times fainter than in non-expressing clones (Figure 7, B-lanes). Since both types of clones contained comparable numbers of proviruses, the results indicate that a significant number of U3pgklxtneoMCCre proviruses had deleted the sequences flanked by loxP.

EXAMPLE 3

To improve the efficiency of self-deletion of proviruses expressing Cre, two additional constructs were made using alternative promoters and selectable marker genes.

Plasmids

pggSVCreU31x was constructed by sequentially inserting the Cre region of MCCre, the SV40 promoter of pBABEpuro and the loxP-fragment of pGEM30 which includes a downstream HindIII site, as blunt ended fragments into the BamHI, XhoI and NheI sites of pGgU3en(-), respectively.
pggSVCreU31xpgkpuro and pggSVCreU31xSVpuro were obtained by ligating blunt ended pgk-puro- or SV40-puro expression cassettes into the HindIII site of loxP.

Cells and viruses

as described for examples 1 an 2.

DNA Hybridizations and PCR

DNA hybridizations were performed with ^{32}P-labeled probes as previously described (21) Southern blots were scanned with a PhosphoImager (Molecular Dynamics) and analyzed with ImageQuantNT software (Molecular Dynamics). For PCR-assays, 150 ng of genomic DNAs were amplified for 40 cycles (94^{o}C 30", 60^{o}C 1', 72^{o}C 2') using the Cre-specific primers 5'-TTAGCTAGCATGCCCAAGAAGAAGAAG-3' and 5'-GGAGCTAGCCTAATCGCCATCTTCCAG-3'. Control PCRs were performed under the same condition except for using actin-primers as previously described (25).

Results

The vectors -pggSVCreU31xpgkpuro- and -pggSVCreU31xSVpuro- (Figure 8) were transfected into BOSC23 cells to obtain infectious virus. Recovered viruses were used to infect NIH3T3 and puromycin resistant clones were isolated after 7 days of selection.

Genomic DNA of puromycin-resistant clones obtained with either construct was analyzed by Southern blotting. To identify recombination of U31xpgkpuroSVCre proviruses, the DNA was digested with EcoR1, a restriction endonuclease that cleaves both proviruses in front of each promoter (Figure 9A, 10A). When hybridized to pgk, non-recombined proviruses maintain all sequences flanked by loxP, and as a result, generate a constant band of 3.1 kb (Figure 9B, lanes 5, 9, 12-14). This band should disappear from Cre-expressing clones (Figure 9A). Accordingly, the band was lost from 9 out of 14 U31xpgkpuroSVCre-expressing clones, indicating that most proviruses have recombined (Figure 9B, lanes 1-4, 6-8, 10). Additional bands of varying sizes represent the fragments extending from the EcoR1 sites of the 3'LTR to sites in the flanking cellular DNA (Figure 9B). However all clones, including those that appeared recombined on Southern blots, generated Cre-specific amplification products when analyzed by PCR (Figure 9C). Although recombined clones generated 10 to 50 fold less amplification product (Figure 9C), the results suggest that the recombinase does not uniformly achieve the threshold levels required for recombination.

Significantly higher recombination efficiencies were obtained with U31xSVpuroSVCre. As shown in Figure 10B, none of the clones generated a 3 kb restriction fragment when hybridized to SV40, and only variable bands consistent with proviral deletion were seen. To confirm recombination, the DNA was digested with HindIII, which cleaves upstream of

puromycin (Figure 10A). As expected, only bands of varying sizes were obtained when hybridized to SV40- or, alternatively, to puromycin probes. Moreover, hybridization patterns were unique, reflecting fragments extending from the HindIII site of the 3'LTR, to sites in the 5' and 3' flanking cellular DNA, respectively (Figure 10B and data not shown). When analyzed by PCR, only 3 out of 12 clones generated faint Cre-specific amplification products (Figure 10C).

The vector system according to the invention can be used for preparing a pharmaceutical composition which contains the conventional carriers and/or diluents. Depending on

the disease to be treated the vector system contains in the additional transcription unit the gene useful for curing the disease.

REFERENCES

1. Akgün, E., Ziegler, M. & Grez, M. (1991) *J Virol* **65**, 382-388.

2. Challita, P. M. & Kohn, D. B. (1994) *Proc. Natl. Acad. Sci. USA* **91**, 2267-2571.

3. Grez, M., Akgün, E., Hiberg, F. & Ostertag, W. (1990) *Proc Natl Acad Sci USA* **87**, 9202-9206.

4. Nusse, R. (1986) *Trends Genet.* **2**, 244-247.

5. Nusse, R. (1991) *Curr. Top. Microbiol. Immunol.* **171**, 44-65.

6. Donahue, R., Kessler, S., Bodine, D., McDonagh, K., Dunbar, C., Goodman, S., Agricola, B., Byrne, E., Raffeld, M., Moen, R., Bacher, J., Zsebo, K. & Nienhuis, A. (1992) *J. Exp. Med.* **176**, 1125-1135.

7. Amariglio, N. & Rechavi, G. (1993) *Environ. Mol. Mutagen.* **21**, 212-218.

8. Cohen, M. & Larsson, E. (1988) *BioEssays* **9**, 191-196.

9. McDonald, J. F. (1993) *Curr. Opin. Genet. Dev.* **3**, 855-864.

10. Nienhuis, A. W., Walsh, C. E. & Liu, J. (1993) in *Viruses as therapeutic gene transfer vectors*, ed.^eds. Young, N. S. (Marcel Dekker, New York), pp. 353-414.

11. Gu, H., Zou, Y. & Rajewsky, K. (1993) *Cell* **73**, 1155-1164.

12. Sauer, B. & Henderson, N. (1988) *Proc. Natl. Acad. Sci. USA* **85**, 5166-5170.

13. Sternberg, N. & Hamilton, D. (1981) *J. Mol. Biol.* **150**, 467-486.

14. Broach, J. R. & Hicks, J. B. (1980) *Cell* **21**, 501-508.

15. Golic, K. G. & Lindquist, S. (1989) *Cell* **59**, 499-509.

16. O'Gorman, S., Fox, D. T. & Wahl, G. M. (1991) *Science* **251**, 13511355.

17. Adra, C. N., Boer, P. H. & McBurney, M. W. (1987) *Gene* **60**, 65-74.

18. von Melchner, H., Reddy, S. & Ruley, H. E. (1990) *Proc Natl Acad Sci USA* **87**, 3733-3737.

19. Morgenstern, J. P. & Land, H. (1990) *Nucl. Acids. Res.* **18,** 3587-3596.

20. Chang, W., Hubbard, C., Friedel, C. & Ruley, H. (1993) *Virology* **193,** 737-747.

21. von Melchner, H., DeGregori, J. V., Rayburn, H., Reddy, S., Friedel, C. & Ruley, H. E. (1992) *Genes & Dev* **6,** 919-927.

22. Pear, W. S., Nolan, G. P., Scott, M. L. & Baltimore, D. (1993) *Proc. Natl. Acad. Sci. USA* **90,** 8392-8396.

23. Reddy, S., Rayburn, H., von Melchner, H. & Ruley, H. E. (1992) *Proc. Natl. Acad. Sci. USA* **89,** 6721-6725.

24. von Melchner, H. & Ruley, H. E. (1995) in *Gene entrapment*, ed.^eds. Farzaneh, F. & Cooper, D. N. (Bios Scientific Publishers, Oxford), pp. 109-129.

25. Schmitt, R., Bruyns, E. & Snodgrass, H. (1991) *Genes & Dev* **5,** 728-740.

26 Morgan, R. A & Anderson W.F. (1993) Annu. Rev. Biochem. 62:191-217.

CLAIMS

1. Vector system useful for gene therapy comprising at least one coding sequence for a site-specific recombinase and at least one target sequence being specifically recognized by said recombinase.

2. Vector system as claimed in claim 1 wherein the recombinase is Cre and the target sequence is loxP.

3. Vector system as claimed in claim 1 wherein the recombinase is Flp and the target sequence is frt.

4. Vector system as claimed in claims 1-3 comprising an additional transcriptional unit.

5. Vector system as claimed in any of claims 1-4 which is a retroviral DNA or RNA.

6. Vector system as claimed in claim 5 comprising the target sequence in the U3 and/or U5 region.

7. Vector system as claimed in claims 5 or 6 wherein the additional transcriptional unit is in the U3 and/or U5 region.

8. Vector system as claimed in claims 1-7 further comprising a viral promoter and/or enhancer.

9. Mammalian cell containing at least one target sequence as in claim 1, at least a foreign transcriptional unit as in claim 4 and essentially no other vector sequences of the vector system according to claim 1 in its genome.

10. Mammalian cell as in claim 9 wherein the target sequence is loxP.

11. Mammalian cell as in claim 9 wherein the target sequence is frt.

12. Process to transduce a cDNA sequence into a mammalian genome, involving the following steps:
- construction of retroviral vectors as in claims 1-8 by essentially standard procedures
- transfection of producer cells
- selection of virus producing clones
- infection of mammalian cells
- expression of a site specific recombinase
- recombination between two or more specific target sequences involving the simultaneous deletion of the recombinase expression cassette and most proviral sequences.

13. Process as in claim 12 involving the following steps:
- construction of retroviral vectors as in claims 1-8 by essentially standard procedures
- infection of mammalian cells
- expression of recombinase
- recombination between two or more specific target sequences involving the simultaneous deletion of the recombinase expression cassette and most proviral sequences.

14. A process for incorporating a nucleic acid into the genome of a mammalian cell comprising the step of introducing a vector system according to any of claims 1-8 into the mammalian cell.

15. Use of a vector according to any of claims 1-8 for incorporating a nucleic acid into the genome of a mammalian cell.

16. Vector system according to any of claims 1 to 8 as a medicament.

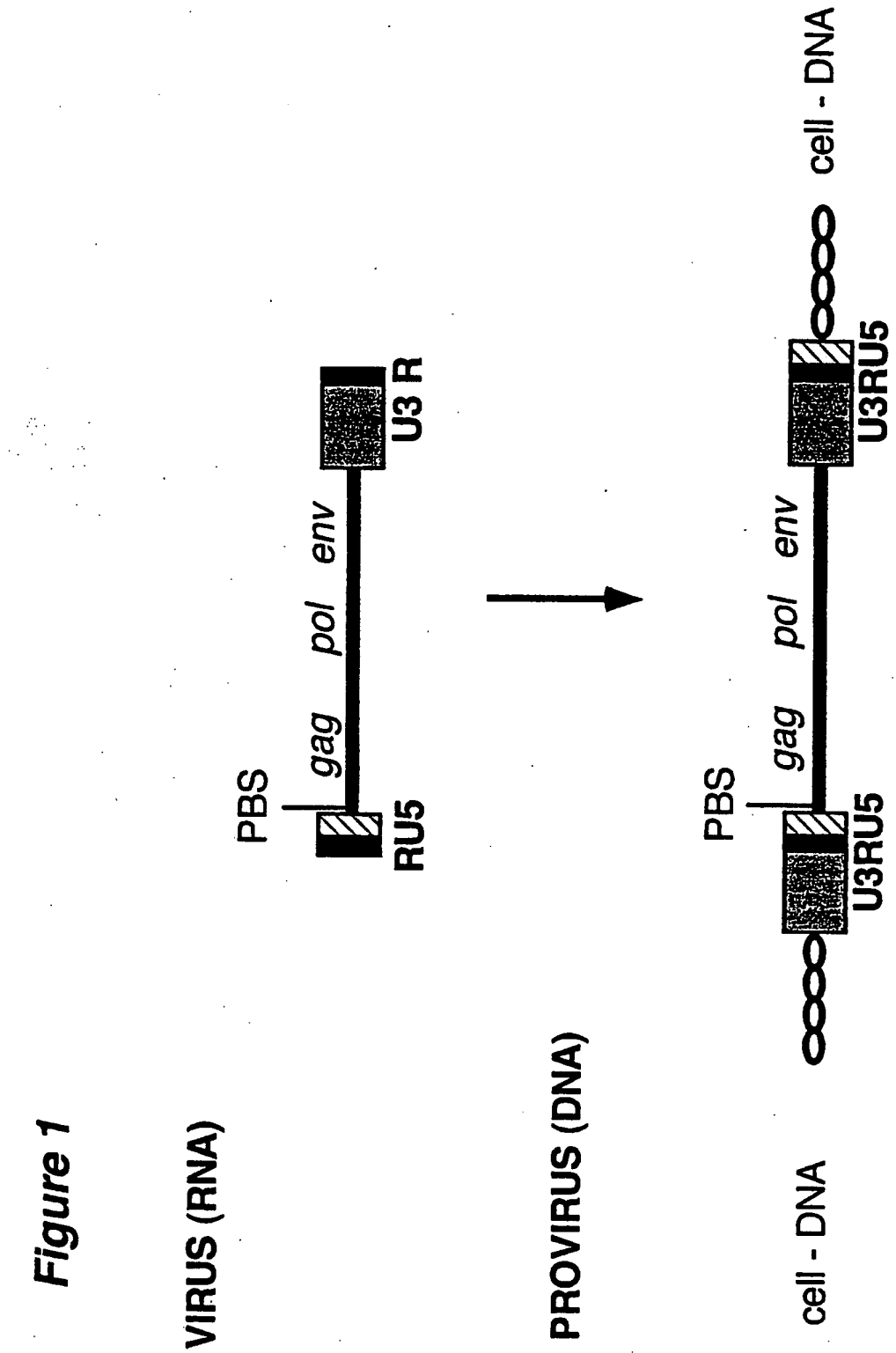

VIRUS (RNA)
PBS
gag
pol
env
U3 R
RU5
PROVIRUS (DNA)
cell - DNA
PBS
gag
pol
env
U3RU5
U3RU5
cell - DNA

Figure 1

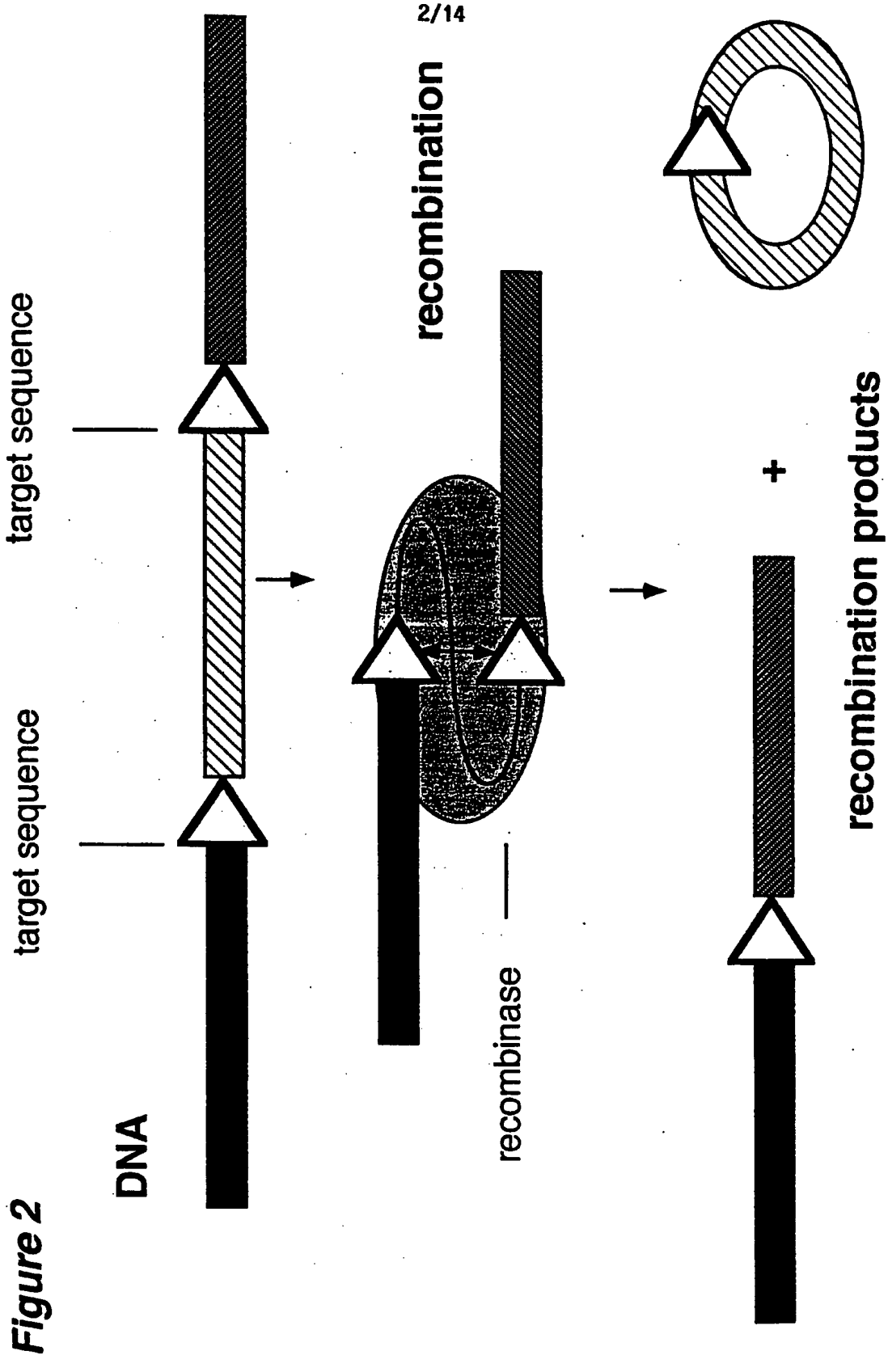
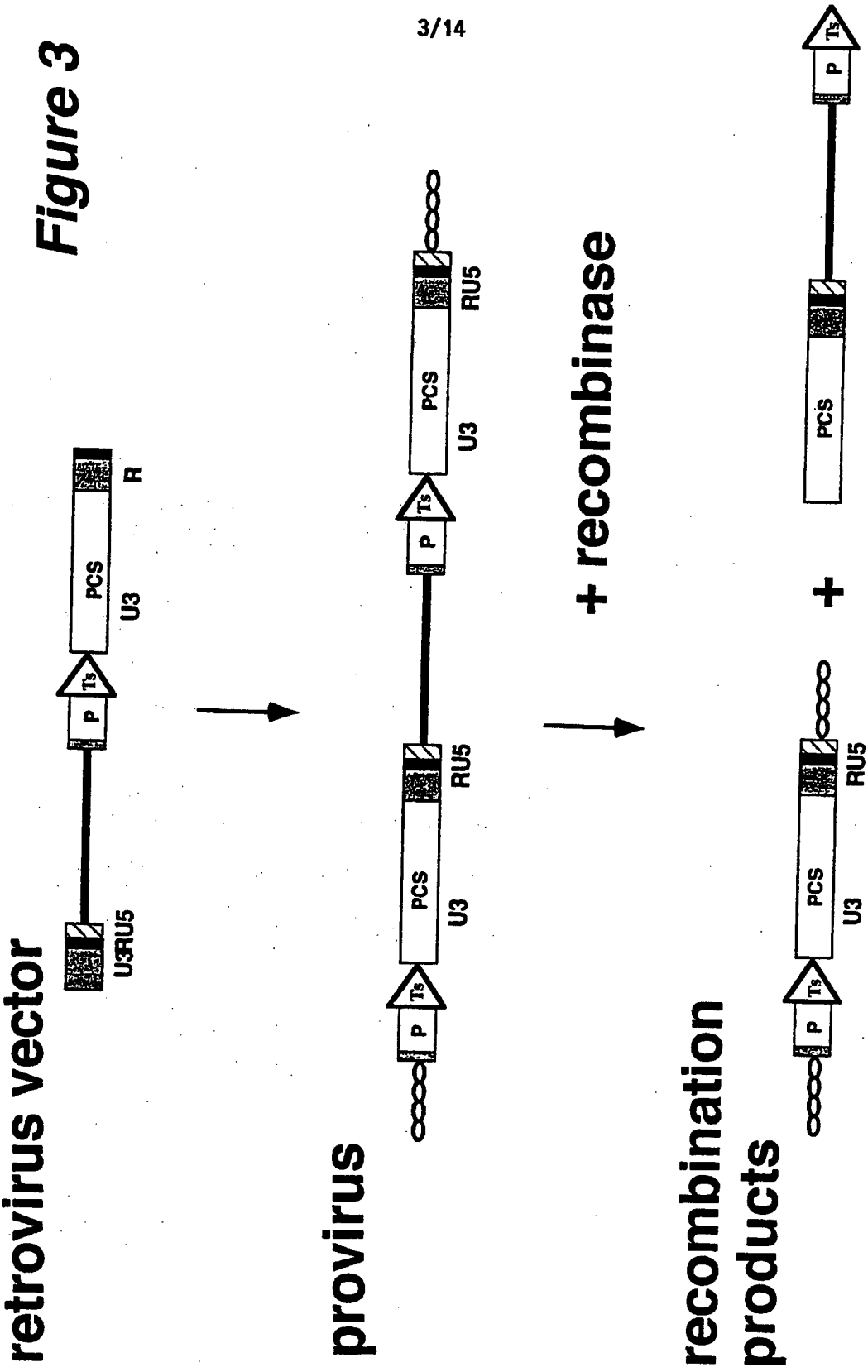

target sequence
target sequence
DNA
recombination
recombinase
+
recombination products

Figure 2

Figure 3


retrovirus vector
U3RU5
PCS
U3
R
Ts
P
provirus
RU5
+ recombinase
recombination products
+